Metric Spaces

let M be set of elements & let :

(metric) $d : M \times M \longrightarrow \mathbb{R}_0^+$ distance

the pair (M, d) is a metric space.

Propirties:

① $d(x,y) = d(y,x)$

② $d(x,y) \geq 0$ $d(x,y) = 0 \Leftrightarrow x = y$

③ $d(x,y) \leq d(x,z) + d(z,y)$

$\forall\, x, y, z \in M$

example

$(V, \|.\|) \rightarrow$ metric spce with normed matric

$d(x,y) = \|x - y\| \geq 0$ $= 0 \Leftrightarrow x = y$

$d(x,y) = \|x - y\| = \|y - x\| = d(y,x)$

$d(x,z) = \|x - z\| = \|x - y + y - z\|$
$\leq \|x - y\| + \|y - z\|$

$$\boxed{d(x,z) \leq d(x,y) + d(y,z)}$$

energy norm

$$\|x\|_A = x^T A x$$

Topology in metric spaces:

let (M, d) be a metric space

① for $x \in M$ $r > 0$ an open Ball

$$B_r(x) = \{ y \in M \mid d(x,y) < r \}$$

② let $U \subseteq M$ is called neighborhood of $x \in U \Rightarrow \exists \varepsilon > 0$ s.t. $B_\varepsilon(x) \subseteq U$

Open & closed Sets:

* let $U \subseteq M$ U is an open set if U is a neighborhood $\forall x \in U$

$$\forall x \in U, \exists \varepsilon > 0 \quad \text{s.t} \quad B_\varepsilon(x) \subset U$$

let $U \subseteq M$ U is closed set if the $M - U$ is open
or $U^c = M - C$ U^c is open.

Remarks:

let (M, d) be a metric space.

① T is set of all possible opensets $\in M$ then T is Topology in M

② M is open & closed cuz $\phi = M^c$ open/closed

③ if U, V are open then $U \cup V$ / $U \cap V$ open
closed closed

④ for U to be open/closed it depends on M and d.

For example

let $(\mathbb{R}, |.|)$
$U = [0, 1]$ is not open/closed
cuz on boundries we can't get $B_\varepsilon(1) \subset U$

let $([0, 1], |.|)$
the $[0, 1]$ is open & closed

⑤ $\{x\} \in M$ is closed

⑥ $B_r(x) = \{y \in M \mid d(x, y) < r\}$ open ball

$K_r(x) = \{y \in M \mid d(x, y) \leq r\}$ closed ball

⑦ $(\mathbb{R}, |.|)$ $[a, b]$ closed (a, b) open
$(a, b]$ neither open nor closed.

Boundries:

let (M,d) mertric space. let $A \subset M$ $x \in M$ is called Boundry if $\forall \varepsilon > 0$ $B_\varepsilon(x)$ contain element of A & M/A.

∂A set of Boundries of A

$int(A) = A - \partial A$

$\bar{A} = A \cup \partial A$

Remarks:

$int(A)$ open, $\bar{A}$ closed, ∂A closed

$int(A) \subset A \subset \bar{A}$

$\partial A = \bar{A} - int(A)$

A closed $\Leftrightarrow \bar{A} = A$ (contains its boundries)

$A = (a,b]$ not open / not closed $\partial A = \{a,b\}$ Closed $\bar{A} = [a,b]$ closed

$int(A) = (a,b)$ open

$\forall x \in A$ $\exists \varepsilon > 0$ s.t. $B_\varepsilon(x) \subset A$ (open set)

Convergence & Compactness

Convergence:

let (M,d) a sequence $(x_k)_{k\in\mathbb{N}}$ is called

Convergent to a limit $a \in M$

if $\forall \varepsilon > 0$ $\exists N \in \mathbb{N}$ s.t. $\forall k \geq N$ $x_k \in B_\varepsilon(a)$

$x_k \xrightarrow{k\to\infty} a$ $\qquad \lim_{k\to\infty} x_k = a$

x_k is convergent $\Leftrightarrow \lim_{k\to\infty} d(a, x_k) = 0$

$\Leftrightarrow \forall \varepsilon > 0,\ \exists N \in \mathbb{N}$ s.t $\forall k \geq N$

$x_k \in B_\varepsilon(a)$

or $d(a, x_k) < \varepsilon$

* if x_k is a convergent sequence then all subsequences are convergent also.

A is closed $\Leftrightarrow \forall x_k \xrightarrow{k\to\infty} a \in A$

example $x_k = \frac{1}{k}$ $\quad k \in \mathbb{N}$ $\quad$ is convergent seq

$\lim_{k\to\infty} x_k = 0$ $\quad$ converges to 0

Cauchy sequence:

let (M, d) $\quad x_k$ is Cauchy seq

$\forall \varepsilon > 0 \ \exists N \in \mathbb{N}$ s.t $\forall n, m \geq N \ d(x_n, x_m) < \varepsilon$

Complete

if all Cauchy sequences converges then it's complet.

* every convergent seq $\rightarrow$ cauchy seq.

Eample)

* $(\mathbb{R}, |\cdot|)$ is complet.

* $(\mathbb{Q}, |\cdot|)$ is not complet.

$x_1 = 1 \qquad x_{k+1} = \frac{x_k}{2} + \frac{1}{x_k} \longrightarrow \sqrt{2} \notin \mathbb{Q}$

* $(V, \|\cdot\|)$ if complete is called Banach space

* $(V, \langle \cdot, \cdot \rangle)$ if complete is called Hilbert space.

ex $(\mathbb{R}^n, \|\cdot\|)$ Banach space

$(\mathbb{R}^n, \|\cdot\|_2) \rightarrow (\mathbb{R}^n, \sqrt{\langle \cdot, \cdot \rangle})$ Hilbert space.

Bounded Set

let (M,d) A bounded $\Leftrightarrow \exists\, \varepsilon > 0,\ x \in A,\ A \subset B_{\varepsilon}(x)$

Compactness

(M,d) is compact if $\forall\, x_k$ has convergent subsequence.

$U \subset M$ is compact if $\forall\, x_k$ has covergent subseq to $a \in U$.

Theorem:

U is compact $\Longleftrightarrow$ U is closed & bounded

exam

$A = [a, b]$ in $(\mathbb{R}, |.|)$

Closed

Bounden let $x = \frac{a+b}{2} \in A$

let $\varepsilon = a+b$ $\quad A \subset B_{\varepsilon}(x)$

A is Compact set.

Continuity:

f is continous on $a \Leftrightarrow \lim_{x \to a} f(x) = f(a)$

f is continous on $V \Leftrightarrow \forall a \in V \quad \lim_{x \to a} f(x) = f(a)$

Properties:

let f, g are continous:

linear Combination are also continous.

if $f_i := M \to \mathbb{R}^n \quad f(x) = f_1 + f_2 + \cdots f_n(x)$

if all $\forall i \in \{1, \ldots n\}$ is continous then f is continous.

ε - δ Continouity

f is continuouse in $a \Leftrightarrow$

$\forall \varepsilon > 0, \exists \delta > 0 \Rightarrow d(x, a) < \delta$ &

$d(f(x), f(a)) < \varepsilon$

Liner - normed Continuity:

let $(M, \|\cdot\|_m)$ $(W, \|\cdot\|_w)$ $f: M \to W$

f is continous $\Leftrightarrow \exists C > 0$ s.t. $\|f(x)\|_w \leq C\|x\|_m$ $\forall x \in M$

Continuity & Topology:

let (M,d) (N,d) $f: M \longrightarrow N$

f is continous $\Leftrightarrow \forall$ open $U \subseteq N \Rightarrow f^{-1}(U) \subseteq M$ open

$\Leftrightarrow \forall$ closed $U \subseteq N \Rightarrow f^{-1}(U) \subseteq M$ closed

$$f^{-1}(U) = \{x \in M \mid f(x) \in U\}$$

like if $f: M \rightarrow R$ continous $c \in M$

$\{f(x) > c\}$ $\{f(x) < c\}$ open

$f(x) \geq c$ or $\leq c$ or $= c$ closed

$\{f(x) = c\}$ level set always closed.

<u>Th:</u> let (M,d) $f: M \rightarrow R$

f is continous & M is compact $\longrightarrow$

then $f(M)$ is compact $\quad Im(M)$

Min/Max Th.

let (M,d) $f: M \rightarrow R$ is continous &

M is compact. the $f(M)$ is bounded

and (Existance of minimizer) $\exists$, $p, q \in M$

$$f(p) = \min_{x \in M} f(x) \qquad f(q) = \max_{x \in M} f(x)$$

Diffrentiation

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

properties

$(f+g)'(x) = f'(x) + g'(x)$

$(fg)'(x) = f'(x)g(x) + g'(x)f(x)$

$\left(\frac{f}{g}\right)'(x) = \frac{f'g - g'f}{g^2}$

chain rule:

$(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$

Partial derivatives (directional derivative)

let $f: R^n \longrightarrow R$

we can find derivative for each var in R^n

exam: $f: R^2 \longrightarrow R$

$f(x,y) = x + y$

$\frac{\partial f}{\partial x} = 1 \qquad \frac{\partial f}{\partial y} = 1 \Rightarrow$ Partial derivative

$\partial_i f(x)$ i-th directional derivative

if $\partial_i f(x)$ is exist $\forall i \in \{1, \dots, n\}$
$\forall x \in$ Domain of f then

f is Partially diffrentionable.

if $\forall i \in \{1, \dots, n\}$ $\partial_i f(x)$ is Continous then

f is Continously diffrentiable.

ex: $f := \mathbb{R}^2 \to \mathbb{R}$

$f(x, y) = x^2 \sin(y)$

$\partial_1 f = 2x \sin(y)$ $\quad\quad$ $\partial_2 f = x^2 \cos(y)$

$\partial_2 \partial_1 f = 2x \cos(y)$ $\quad\quad$ $\partial_1 \partial_2 f = 2x \cos(y)$

if f is twice continously diffrentiable then

$$\boxed{\partial_j \partial_i f = \partial_i \partial_j f}$$

C^K is a set of Function

$C^0 = \{ f \mid f \text{ is Continous} \}$
$C^1 = \{ f \mid f \text{ is Continously diffrentiable} \}$
$C^2 = \{ f \mid f \text{ is twic continously differentiable} \}$
$\vdots$
$C^0 \supset C^1 \supset C^2 \supset \dots C^\infty$

Gradient:

$$\nabla f = \text{vector}(\partial_1 f \; - - - \; \partial_n f)$$

Example

$$f : \mathbb{R}^n \longrightarrow n \qquad f(x) = \|x\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$$

$$\partial_i f = \frac{1}{\cancel{2}} \cdot \frac{1}{\sum_{i=1}^{n} x_i^2} \cdot \cancel{2} x_i$$

$$\partial_i f = \frac{x_i}{\|x\|_2} \Rightarrow \nabla f = \left(\frac{x_1}{\|x\|_2} \quad \frac{x_2}{\|x\|_2} \; ------ \; \frac{x_n}{\|x\|_2}\right)$$

Laplacian:

$$\Delta f = \sum_{i=1}^{n} \partial_i^2 f(x)$$

$$f(x,y) = x^2 + 3y$$

$$\partial_1 f = 2x \qquad \partial_1^2 = 2$$

$$\partial_2 f = 3 \qquad \partial_2^2 = 0$$

$$\Delta f = 2 + 0 = \boxed{2}$$

Hessian: always symmetric

it's a matrix $\in \mathbb{R}^{n \times n}$

$$H_f = \begin{pmatrix} \partial_1^2 f & \partial_2 \partial_1 f & ---- & \partial_n \partial_1 f \\ \vdots & & & \vdots \\ \partial_1 \partial_n f & ----- & & \partial_1 \partial_n f \end{pmatrix}$$

$f(x,y) = x^2 + y^2$ | $f : \mathbb{R}^n \to \mathbb{R}$

$\partial_1 f = 2x$ $\partial_2 = 2y$

$\nabla f = (2x \quad 2y)$

$H_f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ eigenvalues are > 0 so H_f is positive defient.

$\begin{pmatrix} 2x \\ 2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \to$ minimizer

Total derivative

$$\lim_{h \to 0} \frac{\| f(x+h) - f(x) - T(h) \|}{\|h\|} = 0$$

$T(h)$ is total derivatie or $Df(x)$

the norm doesn't matter.

Jaccobian:

let $f: R^n \longrightarrow R^m$

$$Df(x) = J_f(x) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix}$$

$$= \begin{pmatrix} \nabla f_1 \\ \vdots \\ \nabla f_m \end{pmatrix}$$

if $f_1 \, - - \, f_m$ are continously diffrentiable
then $J_f = Df$ = totally diffrentiable.

Chain Rule:

$$J_{g \circ f}(x) = J_g(f(x)) \cdot J_f(x)$$

example:

$f: R \rightarrow R^m \qquad g: R^n \rightarrow R$

$$g \circ f(x) = g(f(x)) = \nabla g \circ f(x) = \nabla g(f(x)) \cdot f'(x)$$

Sub gradient:

Some function are not diffrientiable like $f(x) = |x|$ not diff at $x = 0$

let $f := \mathbb{R}^n \to \mathbb{R}$ a vector $g \in \mathbb{R}^n$ is called subgradient if $\forall z \in \mathbb{R}^n$

$$f(z) \geq f(x) + g^T (z - x)$$

set of subgradient at x is called subdiffrential.

example:

let $f : \mathbb{R} \to \mathbb{R}$ $\quad f(z) = |z|$

$\partial f \; \forall x > 0 = 1 \qquad \partial f \; \forall x < 0 = -1$

at $x = 0 \Rightarrow \quad f(z) \geq f(0) + g \cdot (z - 0)$

$$f(z) \geq gz \Rightarrow |z| \geq gz$$

this is true for $g \in [-1, 1]$

$$g \leq \frac{|z|}{z} \Rightarrow g \leq \frac{z}{z} \quad \text{or} \quad g \geq \frac{-z}{z}$$

$$-1 \leq g \leq 1 \quad \boxed{\in [-1, 1]}$$

Properties of subgradient,

1) $\partial f(x)$ is always closed / convex.

2) All convex & continous Functions are subdiffrientabl

3) if f is Convex & diffrenab $\partial f(x) = \{\nabla f(x)\}$

4) if f is convex and $\partial f(x) = \{g\}$
then f is diffrentiable & $\nabla f(x) = \{g\}$

Tylor expansion:

it's used to approximat a Function arond a valeu x using polynomical

$$f(x+h) = \sum_{k=0}^{\infty} \frac{f^k(x)}{k!} \cdot h^k$$

$f: \mathbb{R} \to \mathbb{R}$ x is constan't k is the numbr the expansion order. h is the approximated ver.

So $\Rightarrow$ $$f(x+h) = \sum_{k=0}^{m} \frac{f^k(x)}{k!} h^k + rest(h)$$

$\lim_{h \to 0} rest(h) = 0$ | m is order of expansion

example $f(x) = \sin(x)$ find second T.E

$$f(x+h) = \sum_{k=0}^{2} \frac{f^k(x)}{k!} h^k + rest(h) \quad \text{arond } x=0$$

$k=0 = \sin(0) = 0$

$k=1 = \frac{\cos(0)}{1!} \cdot h^1 = h \qquad = h - \frac{h^3}{6} \cdots + rest(h)$

$k=2 = \frac{-\sin(0)}{2!} \cdot h^2 = 0$

$k=3 = \frac{-\cos(0)}{3!} h^3 = \frac{-h^3}{6}$

Multi index

$\alpha = \{\alpha_1 \text{ ---- } \alpha_n\}$
$= \{1, 2, 3, 4\}$

operation

1) $|\alpha| = \sum_{i=1}^{n} \alpha_i = 10$

$$\begin{array}{r} 48 \\ 6 \\ \hline 272 \end{array}$$

2) $\alpha! = \prod_{i=1}^{n} \alpha_i! = 1!\ 2!\ 3!\ 4! = 1 \times 2 \times 6 \times 24$
$= 272$

3) $x^\alpha = \prod_{i=1}^{n} x_i^{\alpha_i}$ for $x \in R^n$

4) for $f : R^n \to R$ $\quad f \in C^{|\alpha|}$ then

$$\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \partial_3^{\alpha_3} \text{ --- } \partial_n^{\alpha_n} f$$

$\alpha = \{1, 2\}$

$f = x^3 + y^3$

$\partial^\alpha f = \quad \partial_1^1 \partial_2^2 = \partial_1^1 (6y) = 0$

Multidim Tylor expansion

$$f(x+h) = \sum_{|\alpha|<l} \frac{\partial^\alpha f(x)}{\alpha!} \cdot h^\alpha + r(h)$$

Second Tylor exarsion

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T H_f(x) h + r(h)$$

$$f(x+h) = f(x) + \nabla f^T(x) h + \frac{1}{2} h^T H_f(x) h + r(h)$$

Optimization

$$x^* = \min_{x \in D} f(x)$$

Existance of minimizer:

let (M,d) metric space ; $f := M \to \mathbb{R}$

f is bounded from below. $\forall x \in M$ $f(x) \geq C$

then C is min $f(x)$

to check the exisitance.

① M is compact $:\Rightarrow \exists x_k$ it has convergent subsequence

② f is continouse $\Rightarrow$ at a $\Rightarrow f(a) = \lim_{x \to a} f(x)$

$\boxed{\text{So}}$ Theorm

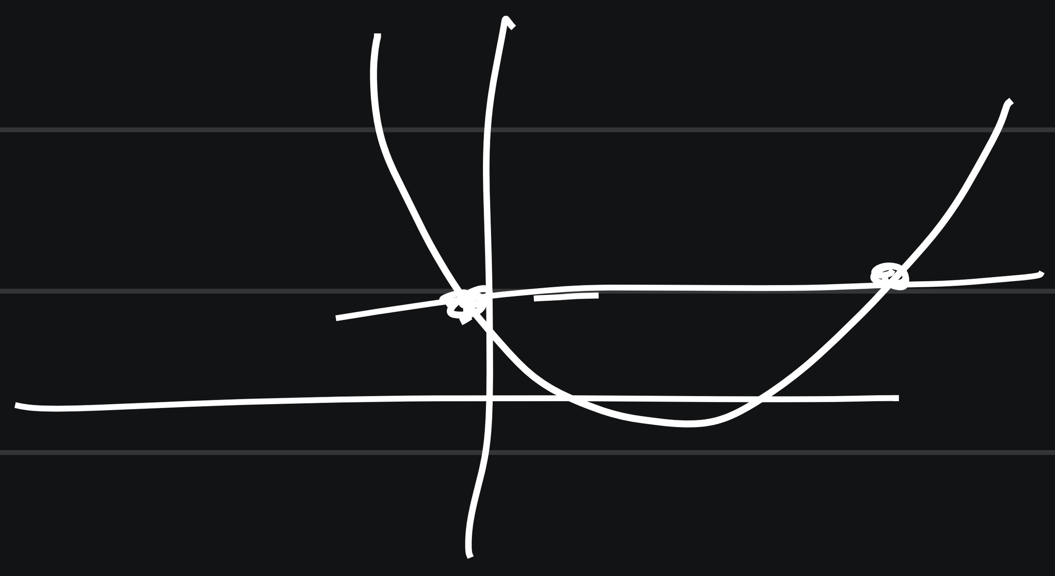
Let (M,d) $U \subset M$ $f := U \to \mathbb{R}$

if U is compact (closed & bonded) and f is contiuen then there is a minimizer $x^* \Rightarrow$

$$f(x^*) = \min_{x \in U} f(x)$$

Uniqueness :

Convex function



let V linear space $A \subset V$ $f : A \to R$

f is convex $\Rightarrow f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$

convex

$<$ strictly convex

all P-norm functions are convex.

function between any two points is under the chord between the two points.

so $f(\underbrace{\lambda x + (1-\lambda) y})$

Point betwen x , y (using weighted average

ex) $x_1 = 1$ $x_2 = 5$

$x_3 = 0.5 \times 1 + 0.5 \times 5 = \frac{1}{2} + \frac{5}{2} = \boxed{3}$

$x_4 = 0.1 \times 1 + 0.9 \times 5 = \frac{1}{10} + \frac{45}{10} = 4.6$

10% of 1 90% of 5

$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$

weighted avg between $f(x)$ & $f(y)$

$\boxed{0 \leq \lambda \leq 1}$

Theorm:

$f \in C^2$

f convex $\iff$ $H_f(x)$ is Positive semi defient

and $H_f(x)$ is positive defient $\Rightarrow$ f is stricktly convex.

Uniqueness

if f is stricktly convex then and f has x^* minimizer then x^* is unique.

To identify a local min/max

let $f: \Omega \to \mathbb{R}$ | $f \in C^2$ $\nabla f(x) = 0$

if $\exists x^* \ni f(x^*) = \min/\max_{x \in \Omega} f(x)$

x^* local min $\rightarrow$ $H_f(x)$ is positive semi defient
local max $\rightarrow$ H_f is negative semi defient

isolated min $\rightarrow$ H_f Positive defient
isolated max $\rightarrow$ H_f negative defient

example

$f : \mathbb{R}^2 \to \mathbb{R}$

① $f(x,y) = x^2 + y^2$

candidate

$\nabla f(x) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$H_f = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ $\forall \lambda$ eigenvalue > 0
so H_f is positive definite

then f is strictly convex then x^* is isolated minima.

Necessary condition to get x^* as minima.

$$\nabla f(x^*) = 0$$

Suffient $\nabla f(x^*) = 0$ & $H_f(x^*) > 0$

$\Downarrow$

f is strictly convex